Controlling Signalized Intersections using Machine Learning

Eriksen, Andreas Berre; Lahrmann, Harry; Larsen, Kim Guldstrand; Taankvist, Jakob Haahr

Published in:
Transportation Research Procedia

DOI (link to publication from Publisher):
10.1016/j.trpro.2020.08.127

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 48 (2020) 987–997

Transportation Research Procedia

www.elsevier.com/locate/procedia

World Conference on Transport Research – WCTR 2019, Mumbai, 26-30 May 2019

Controlling Signalized Intersections using Machine Learning

Andreas Berre Eriksen[a], Harry Lahrmann[b]* Kim Guldstrand Larsen[a] , Jakob Haahr Taankvist[a]

[a] Department of Computer Science, Aalborg University, Selma Lagerlöfs Vej 300, DK-9220 Aalborg East, Denmark
[b]Department of Civil Engineering, Aalborg University, Thomas Maans Vej 23, DK-9220 Aalborg East, Denmark

Abstract

Signalized intersections are the capacity-determining points on roads in cities, and the signal settings are usually based on very primitive algorithms which cause road users to experience a lot of unnecessary delay The work presented in this paper, show the effect of deploying a controller based on the optimization software Uppaal Stratego in four signalized intersections on the same road segment, the controller is fully distributed meaning there is no direct coordination between the intersections. The controller is tested against the controller deployed in the intersections today. The controllers have been tested using the micro simulation program VISSIM. The simulation shows that in comparison with the existing controller, this controller provides a reduction of between 30% and 50% in average delays, queues and number of stops. The fuel consumption and total travel time of cars in the coordinated section are reduced by about 20% in the simulation study. All these reductions are achieved without making the situation worse for the side roads.

Keywords: Signalized Intersections; optimization; Machine Learning; Reinforcement Learning; Model Checking; Uppaal Stratego

1. Introduction

Signalized intersections are necessary to maintain good traffic performance and a high safety level on our roads, but signal systems also generate several inconveniences for the road users and society in general. These include operation and maintenance costs, stops in the traffic, increases in delays and high fuel consumption. The issues caused by signal systems largely depend on the control system used [Lauritzen, 1994]. In Denmark, it is estimated that the annual societal costs of a Signalized intersection amount to EUR 650,000 of which delay related costs

* Corresponding author. Tel.: +45 4056 0375
E-mail address: hsl@civil.aau.dk

10.1016/j.trpro.2020.08.127

represent 69% of this 2012 [Vejdirektoratet, 2012]. Therefore, Signalized intersections represent a great potential for optimizing traffic flows.

A signalized intersection can be either time or traffic controlled. In addition, the system can be independent of other intersections or coordinated with other intersections and the coordination can be simple or advanced – often called adaptive signal management [Gautier, 2001]. The first adaptive signal systems were developed in the late 1970s and have since been further developed and new systems have been developed. SCOOT and SCATS are the most common, and together they cover 80% of the market, but there are also systems such as SPOT and MOTION. Common to the systems is that they are based on historical data or feedback from loops. Depending on the system, changes in program, cycle length or offset time typically take one or more minutes, and in some cases, up to 10 minutes. [Kronborg & Davidsson, 2004; Lauritzen, 1994].

Reliable detection of traffic is important and detection can be carried out using many different detection methods. Induction loops detects traffic in point and has always been the most widely used technology, but in recent years traffic radar has become commonplace. The radar technology introduces the so-called Estimated Time of Arrival (ETA) function, which continuously estimates the arrival time of cars to the stop line based on course, speed and position. Most of today's systems do not use the continuous flow of ETA values from the radar, instead they emulate virtual induction loops and thus translate the continuous flow of ETA values into discrete observations of ETA when cars pass a virtual loop. [Jakobsen, 2015; Kildebogaard, 2015; Traffic Rader | Intersection Management, 2016]. Today's control of signal systems builds on a few simple signal controlling principles. Typical elements are predetermined phase order fixed offset times, fixed maximum cycle times, and fixed maximum and minimum green times. Even no binding regulation applies to these control elements strong traditions seems to influence the choice of values to be used.

The starting point for this project is to use modern optimization techniques to optimize traffic flow in signalized intersections. Others have done this before and examples of this are: [Mousavi, 2017] views traffic in intersections as MDP with Reinforcement Learning and Deep Policy-Gradient and train a deep neural network for adaptive signal control. The resulting controller was evaluated on a small example with respect to different metrics using synthetic traffic load. No comparison to existing controllers were made, nor does the paper evaluate with respect to actual traffic. [Genders, 2016] use Q-learning for off-line training of deep neural networks representation of traffic light control policies for different optimization criteria. The resulting controllers are experimentally evaluation using the simulation tool SUMO on a small synthetic intersection. [El-Tantawy, 2013] presents an extensive comparison of different off-line learning methods (Q-learning, SARSA, TD) for traffic control based on MDP models based on evaluation of different traffic scenarios from Toronto. [Balaji, 2010] applies Q-learning to off-line optimize green timing in an urban arterial road network to reduce total travel time. The method was tested on simulation of traffic in highly congested section of the Central Business District area in Singapore. [Sanchez-Medina, 2010] use genetic algorithms and cellular-automata-based simulators for *off-line* training optimal traffic signal control. The resulting controller was tested on simulation of contested scenarios for "La Almozara" in Saragossa, Spain. [Wei, 2018] build a frame-work, where deep reinforcement learning is used to train a traffic controller off-line, the off-line training is followed by an online learning phase, where the controller chooses it's action according to an ε-greedy strategy combining exploration and exploitation. The controller is tested using real world data and shows an improvement in delay of 19 %

This project is based on [Eriksen, 2017] which use Uppaal Stratego for *online* synthesis of optimal control for a signalized intersection. Uppaal Stratego is an optimization program that combines machine learning and model checking techniques to synthesize at run-time a near-optimal control strategy for signal management [David, 2015]. Importantly, and in contrast to above-mentioned applications of Reinforcement Learning to signal controlling, there is no training phase (say for training a Neural Network). In fact, our method will immediately adapt to any changes in the pattern of traffic, and will allow instantaneous change of optimization criteria. We use as a case a road segment from Hobrovej, Aalborg with four signalized intersections. The aim of the project is to develop a new controller based on real time optimization of the signal setting from continuous radar information, with a minimum of restrictions on phase order and minimum/maximum green times. For evaluation we compare the difference between the existing controller and the new controller.

2. Method

The project is based on a section in the street Hobrovej in Aalborg, Denmark with four signalized intersections, see Figure 1, the intersections at each end of the section are only used to send cars into the simulation model. The route is an important entry road to the city of Aalborg and has an AADT of 20,000-30,000 vehicles [Mastra, 2017]. Three of the intersections are today pre-timed operated and coordinated in a green wave with partial traffic control and the last intersection – named Søndre Skovvej - is fully traffic controlled. The four intersections are optimized in relation to today's standard for the regulation of respectively coordinated intersections with partial traffic control and full traffic-controlled intersections in Denmark today.

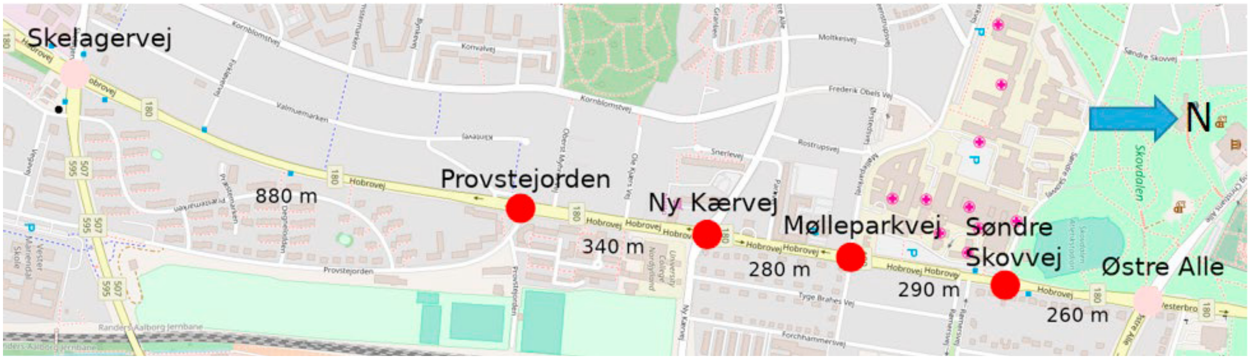

Figure 1 - The six signalized intersections in the street Hobrovej and the distance between the intersections.© OpenStreetMap contributors.

2.1. Traffic modelling

The project uses the VISSIM microsimulation program. This program allows modelling traffic, thus enabling comparing traffic flow with the current controller and the controller developed in this project. VISSIM is used to measure delays, queue lengths, number of stops, fuel consumption and total traveling time on the main road. [PTV VISSIM, 2017].

2.2. Data Collection and Processing

As input to VISSIM we used intersection counts from all six intersections in the morning peak hours from 7:00 to 9:00. The counts were made in 15-minute intervals. Pedestrians and cyclists are not included in this project.

2.3. The Controller

In the paper we have developed a controller, which optimizes the signal setting based on reinforcement learning using continues input information about all cars within 200 meters of the intersections stop line. The control strategy is synthesises using the software Uppaal Stratego [David, 2015]. We used the micro simulation tool VISSIM to generate the input traffic and evaluate the effect of the control strategy. In this we only use a single optimization criterion namely the overall delay in each of the four intersections, we will see later in the results that even though we only optimize for average delay the derived effects on all measures are positive.

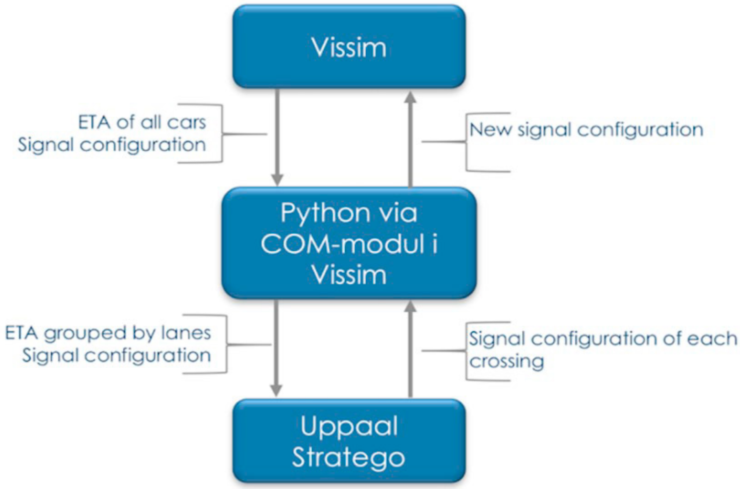

Figure 2 - The communication between Vissim, Python and Uppaal Stratego.

Data equivalent to radar data is extracted from VISSIM through the COM-interface, which is available in the programming language Python, see Figure 2. The data is filtered and partitioned into four subparts one for each intersection. For each intersection we track the ETA (Estimated Time of Arrival) of all cars within 200 meters of the intersection each second. We then run Uppaal Stratego for each intersection with the car information and signal setting that belongs to that intersection. Uppaal Stratego uses this data to build an abstract MDP model from which a near-optimal signal setting for the intersection. is then generated. The signal setting is then read in Python and effectuated in VISSIM. This communication pattern is repeated each second.

The control of each intersection is performed in a distributed manner meaning that there is no communication or coordination between any of the intersection. In Uppaal Stratego we can optimize for any number of parameters, for example queue length, number of stops and delay. In this paper the signal setting is optimized to minimize the overall delay of all cars.

2.4. The Uppaal Stratego Controller

The optimal signal setting is decided *on-the-fly* using the optimization tool Uppaal Stratego: Every second an abstract (MDP) model is initialized using the current traffic situation in the intersection. The information used to make the model are the vehicles detected by the radars, as well as estimates of how well the traffic is moving based on the previous traffic. As an example, we model and optimize a simplified intersection seen in Figure 3 in Uppaal Stratego: This intersection only has two lanes one from East to West and one from South to North. In the current situation, 4 vehicles are waiting from East and 2 from South. Based on previous traffic, the arrivals of vehicles from both East and South are assumed to be exponentially distributed with rates 1, i.e. $r_S = r_E = 1$. In the current situation, the signal setting is green from east to west.

We also have parameters in the model describing how fast the traffic moves away from the intersection. Here the rates of this model are $r_W = -4$ ($r_N = -5$), meaning that when the signal is green in East-West direction (South-North direction) the traffic moves away from the intersection at these rates.

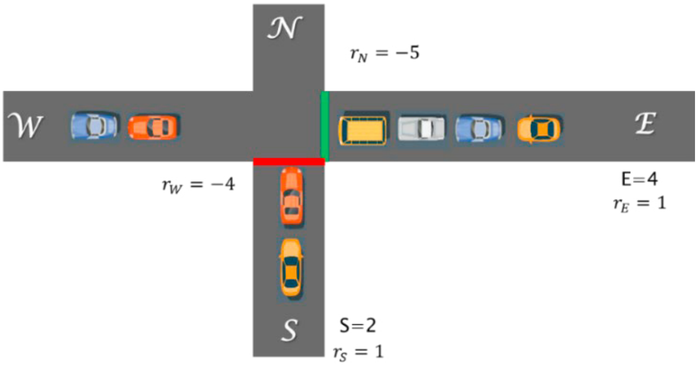

Figure 3: Overview of the example intersection.

The signal setting may be changed every second and the underlying mathematical model is a Markov Decision Process (MDP) [Puterman, 1994]. An MDP is illustrated in Figure 4.

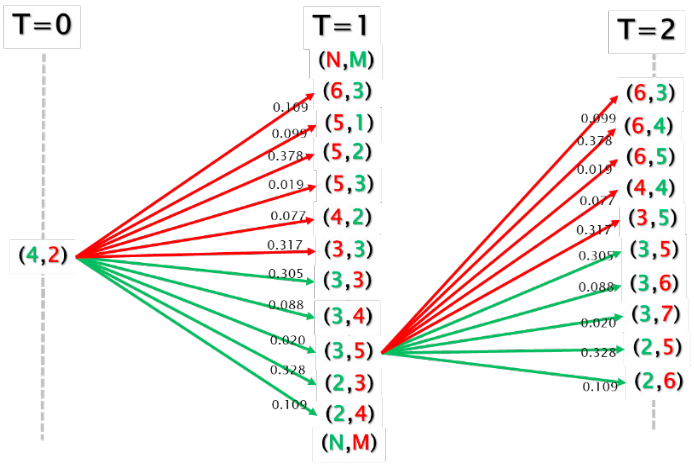

Figure 4 - Example of an MDP. (x,y) indicates the configuration with x cars waiting from East and y cars waiting from South

As we know our current configuration at time $T = i$, we will have a distribution of states we can end up in depending on the choice we make. Our choices in this model is to extend green time or give green to the other direction. Faced with this MDP we can now find an optimal control strategy, which optimizes the objective the model is given. E.g. total delay of all vehicles. To solve the problem, we use Reinforcement Learning which performs an iterative simulation of the MDP-model with the goal to learn the optimal control strategy. We model the choice of phases in Uppaal Stratego using a state machine augmented with timing constraints, expressed using limitations on local clocks (a so-called timed automata). This enables us to guarantee things like minimal green times. In the state machine in Figure 5 x is a local clock, which is reset every time we change the signal setting. The test $x >= 4$ ensures that we only change the setting after four or more seconds. If this test is replaced by the constraint $x == 4$, then we describe a pre-timed operated system where the setting is changed after exactly four seconds. The test $y == 1$ on the local clock y ensures that phase changes may take place every second.

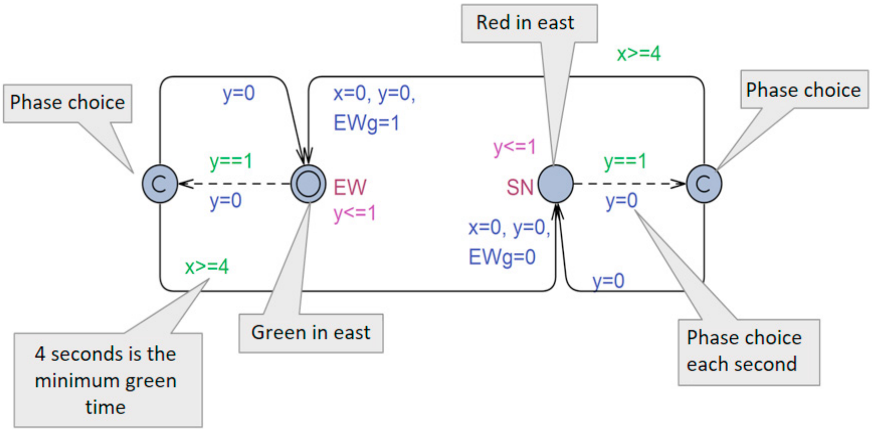

Figure 5 – Uppaal Stratego model of the intersection seen in Figure 3.

In Figure 6, we see a simulation in Uppaal Stratego of a pre-timed operated system over 100 seconds. In the figure we see the two queue lengths as functions of time and the blue curve shows (as 0 or 1) the signal setting. As we see the setting changes after exactly four seconds. An estimate of the accumulated total queue length over 100 seconds is 1411±42.38.

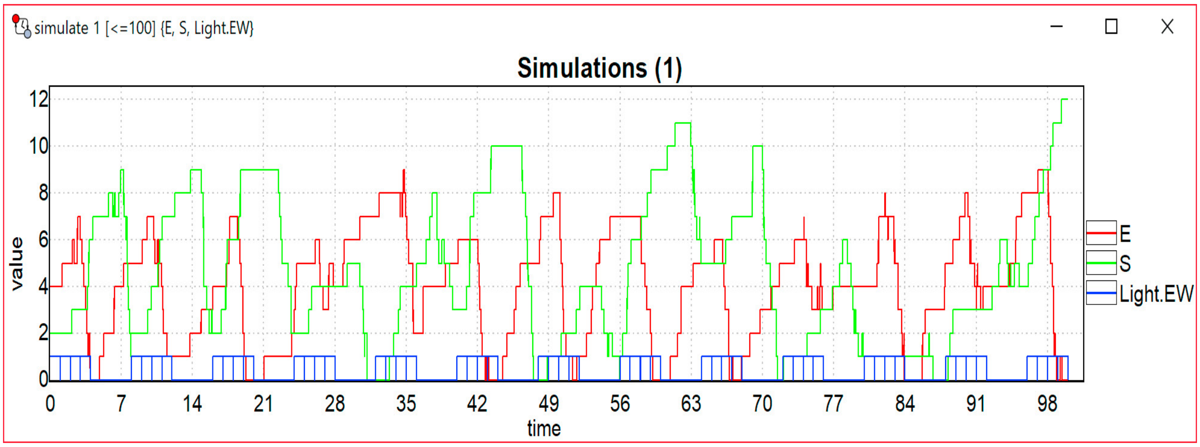

Figure 6 – A simulation of the example intersection using a pre-timed operated system.

For comparison, we show in Figure 7 a simulation of a trained optimal control strategy. As we can see the system is not pre-timed operated, but controlled by the actual traffic in the system. An estimate of the total queue length over 100 seconds is 829±14.57 which means a reduction of 41%.

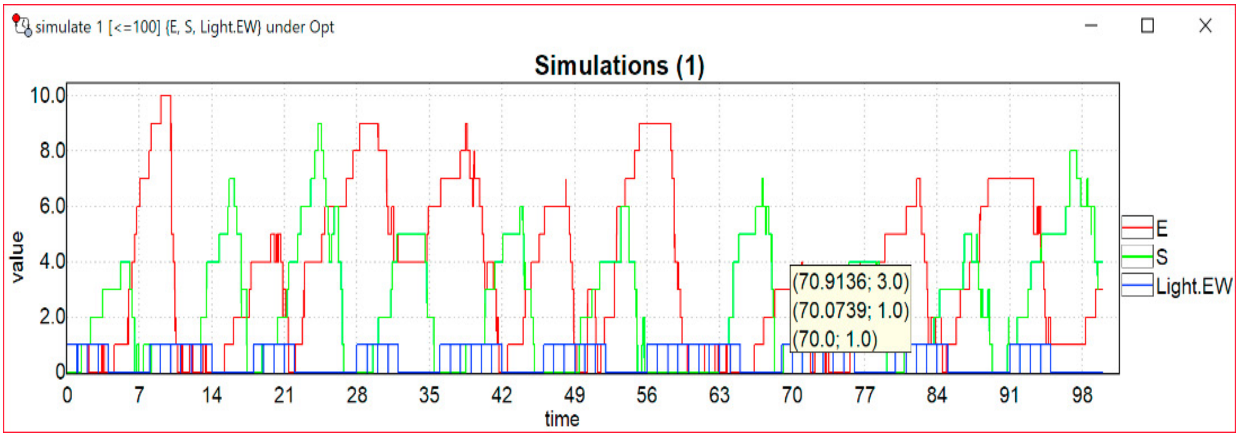

Figure 7 – A simulation of the example intersection using a learned optimal controller.

Above we calculated (learned) the optimal signal setting using Uppaal Stratego based on a model of a snapshot of the intersection, which means we only have information on the vehicles up to 200 meters from the intersection. It is clear that the situation can change significantly in a given horizon (e.g. 20 or 100 seconds). This means that more or less vehicles than expected compared to the assumptions could arrive in the intersection. To combat this problem we introduce a rolling horizon. The rolling horizon gives the opportunity to continuously adapt the model and by that the signal setting to the actual traffic, see Figure 8. This means the decision about the signal setting will be updated once every second

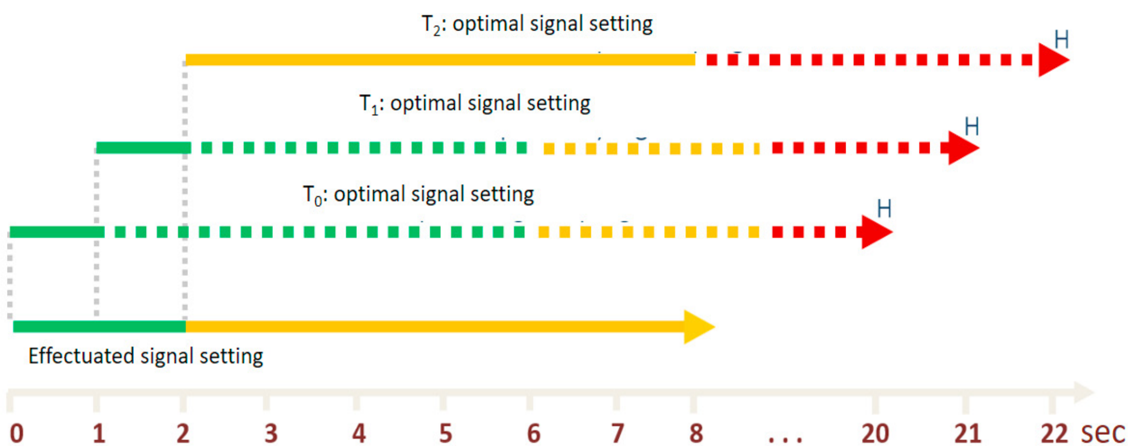

Figure 8 We learn a strategy up to a horizon, we then after a second learn a new strategy using the updated information from the radar.

3. Results

The effect of the new controller is documented by comparing the simulation results from the existing controller with the simulation results from the new controller. The evaluation measures the average delay, mileage, fuel consumption and number of stops made in the individual intersection. In addition, the total travel time at Hobrovej is also measured see Figure 1. Detailed data for the individual traffic flows are presented for the intersection Hobrovej/Ny Kærvej, this is the largest of the intersections.

We see on Figure 9 that the new controller has a significant effect for all the measured parameters in all the four intersections. The average delay is reduced between 27 % and 54 %, the queue lengths are reduced 42-64 %, the number of stops 20-59 % and the fuel consumption 19-28 %. In addition to this the total travel time for the whole segment of Hobrovej through the four intersections is reduced by 29% and 16% for respectively the north and south bound directions.

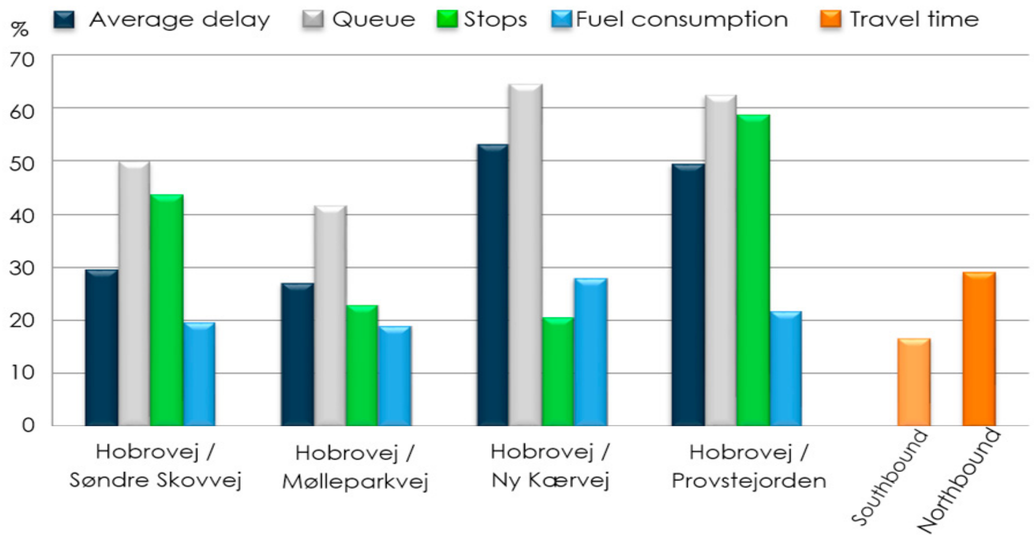

Figure 9 – Summarized results for all intersections. The graph shows the percentwise reduction in the measures using the new controller.

From Table 1 we observe that the controller reduces the average delay for all directions. This shows that the controller can exploit the observed flow of the traffic better than the existing controllers in the four intersections. This means the new controller has reduced the average delay for all streams, and that it does not just moved the delay from the main roads to the side roads.

Table 1 - Average delay in the respective legs with existing controller and new controller, as well as the absolute and relative effects for the intersection of Hobrovej/Ny Kærvej/Vestre Allé.

Traffic flow		Existing controller	New controller	Number	Effect	
From	To	[sec./veh.]	[sec./veh.]	[veh.]	[sec./veh.]	[%]
Hobrovej S	Hobrovej N	14	5	1819	-8	-60
Ny Kærvej	Vestre Allé	30	16	340	-14	-46
Ny Kærvej	Hobrovej S	35	22	174	-13	-36
Vestre Allé	Ny Kærvej	46	17	411	-28	-63
Vestre Allé	Hobrovej S	48	14	139	-34	-71
Vestre Allé	Hobrovej N	50	27	147	-23	-46
Ny Kærvej	Hobrovej N	29	15	333	-15	-50
Hobrovej S	Ny Kærvej	13	5	290	-9	-65
Hobrovej S	Vestre Allé	21	13	159	-8	-37
Hobrovej N	Ny Kærvej	26	18	252	-8	-31
Hobrovej N	Vestre Allé	9	5	95	-5	-50
Hobrovej N	Hobrovej S	9	5	1,337	-5	-49
Weighted average		20	9	5,495	-11	-53

To capture the new controller's performance with respect to the green wave the travel time on the main road is captured for both controllers this is shown in Figures 10 and 11.

The results shows that the highest performance difference is when the congestion is at its highest. This result shows that the new controller is able to constantly avoid being in a state where the intersection is fully congested.

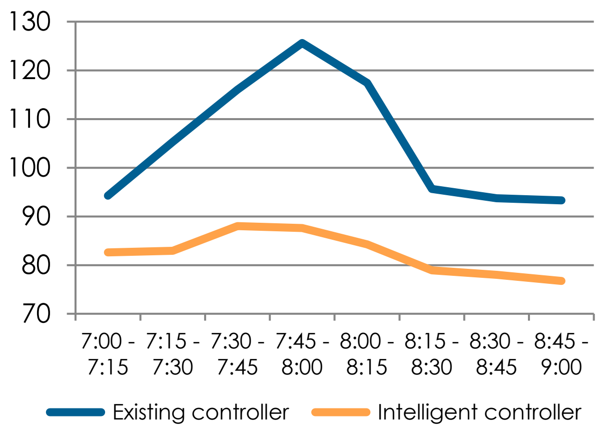

Figure 11 – The northbound travel time

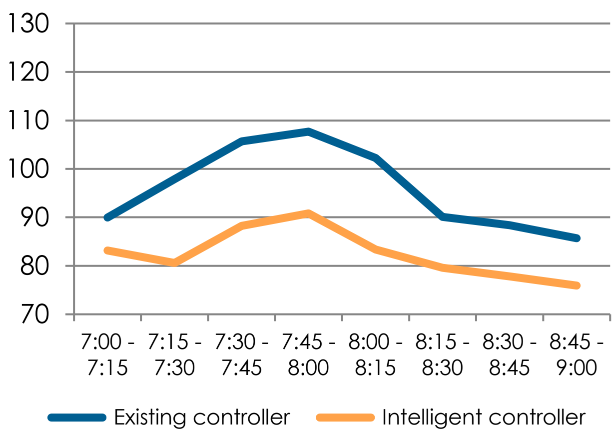

Figure 10 – The southbound travel time

4. Discussion and Conclusion

Through microsimulation with VISSIM, we have demonstrated that a new Uppaal Stratego based controller makes traffic significantly more effective than the existing controller for the metrics average delay, queue length, number of stops, fuel consumption and overall travel time. All parameters were improved in the four intersections and in addition, the total travel time in the previously coordinated route was also significantly improved. The controller reduces the measured parameters for all streams, it does not just move capacity from the side roads to the main direction.

As mentioned earlier, in this project, we only use a single optimization criterion, the overall delay in each of the four intersections. However, with the methods we have used, we can include any other measurable criteria in the optimization - for example a criterion that the queue length in a turning lane must not exceed the length of the turning lane. Another example could be to prioritize the public transport in the optimization, and even include the number of passengers in the bus in the optimization. It is also possible work with dynamic constraints in the optimization, for example, that the maximum red time depends on traffic density so that it rises when traffic rises.

5. Acknowledgments

This work was supported by the Innovation Fund Denmark through the Foundation's funding of DiCyPS – Center for Data-Intensive Cyber-Physical Systems –at Aalborg University.

6. References

1. Balaji, P.G.; German, X.; Srinivasan, D.: 'Urban traffic signal control using reinforcement learning agents', IET Intelligent Transport Systems, 2010, 4, (3), p. 177-188, DOI: 10.1049/iet-its.2009.0096
2. David A, Jensen PG, Larsen KG, Mikučionis M and Taankvist JH 2015, Uppaal Stratego. 21st International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS).
3. Samah El-Tantawy, Baher Abdulhai & Hossam Abdelgawad (2013): Design of Reinforcement Learning Parameters for Seamless Application of Adaptive Traffic Signal Control, Journal of Intelligent Transportation Systems: Technology, Planning, and Operations, DOI: 10.1080/15472450.2013.810991
4. Eriksen AB, Huang C, Kildebogaard J, Lahrmann H, Larsen KG, Muniz M & Taankvist JH, 2017, Uppaal Stratego for Intelligent Traffic Lights, 12th ITS European Congress, Strasbourg, France, 19-22 June 2017.
5. Gautier, E 2001, Adaptive signalsystemer. Dansk Vejtidsskrift (now Trafik & Veje), maj 2001, pp 36-37.
6. Genders, W., & Razavi, S. (2016). Using a Deep Reinforcement Learning Agent for Traffic Signal Control. CoRR, abs/1611.01142.
7. Jakobsen, L. 2015, Intelligent detektering i trafiksignaler. Webinar, ITSDanmark, http://itsdanmark.dk/_files/2015020515.02intelligentdetekteringitrafiksignaler.wmv, Accessed 2. april 2017.
8. Kildebogaard, J. 2015, Trafiksignaler - hvordan udnytter vi udviklingspotentialet? Trafik & Veje, september 2015, side 28-30.
9. Kronborg, P. & Davidsson, F. 2004, Adaptiv styrning av Stockholms trafiksignaler, Kunskapssammanställning och förslag, Movea Trafikkonsult AB.
10. Lauritzen 1994, "Signalregulerede kryds" i Vejtrafik – Trafikteknik & Trafikplanlægning, Lahrmann, H. & Leleur, S., Polyteknisk Forlag. Side 147-180.
11. Mastra 2017,
www.kmastra.dk, Accessed 5 December 2017
12. Puterman M.L.1994, Markov Decision Processes. Wiley and Sons, 1994.
13. Mousavi, Seyed Sajad; Schukat, Michael; Howley, Enda: 'Traffic light control using deep policy-gradient and value-function-based reinforcement learning', IET Intelligent Transport Systems, 2017, 11, (7), p. 417-423, DOI: 10.1049/iet-its.2017.0153
14. PTV VISSIM 2017,
http://vision-traffic.ptvgroup.com/en-us/products/ptv-vissim/, Accessed 5 December 2017.
15. Traffic Rader|Intersection Management, 2016, Smartmicro,
http://www.smartmicro.de/trafficradar/intersection-management/, Accessed 4 April 2017.

16. Vejdirektoratet, 2012. Konsekvenser for trafikanter og samfund – Bedre trafiksignaler, København.
17. Wei H, Zheng G* (equal contribution), Yao H & Li Z, 2018, IntelliLight: A Reinforcement Learning Approach for Intelligent Traffic Light Control, in Proceedings of the 2018 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'18).